nature research

Corresponding author(s): Sihem Cheloufi, Ying Zhang, Tong Zhou, Qi Chen

Last updated by author(s): Feb 15, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The raw small RNA sequencing data were obtained by custom code from BGI, and UCSD IGM Genomics Center. The raw mRNA sequencing data were obtained from Novogene.
Data analysis	Small RNA sequences were annotated using the software SPORTS1.1(v1.1.0). Transcriptome data were annotated using the software kallisto (v0.46.1). The R package DEGseq(v1.40.0) was performed to identify the differentially expressed sncRNAs and the R package edgeR(v3.9) was employed to identify the differentially expressed genes. The R package clusterProfiler(v3.16.1) was used to determine enriched biological process pathways. The FAIME(v2) algorithm was applied to calculate geneset scores based on the rank-weighted gene expression of individual samples. The RNAfold tool in the ViennaRNA (v2.4.14) package was performed for secondary structure prediction. Translation assay data was analyzed by GraphPad Prism (v8.4.3). ImageJ bundled with 64-bit Java 1.8.0_172.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA sequencing datasets have been deposited in the Gene Expression Omnibus (GSE144666). LC-MS/MS data have been deposited in Figshare (https://figshare.com/articles/dataset/_/14033003).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Power analyses were performed to determine the sample size. The power analysis for the RNA-seq experiments were performed by the R package "ssizeRNA" in our pilot RNA-seq datasets. The dispersion level was estimated by the weighted likelihood empirical Bayes method. Three biological replicates were included for each group in the RNA-seq experiments, which would be sufficient to detect >80% of the differentially expressed RNAs with a FDR<0.05 and fold change>1.5. (except two biological repeats were included for mouse spleen and human hESC-naïve samples, which were not included for statistical analyses). We used the R package "pwr" to perform the sample size estimation for the LC-MS/MS experiments and the protein syntheses rate analysis. Three biological replicates were included in the LC-MS/MS experiments, which ensured a statistical power at 1-β of 80% and α of 5%. ~40 samples per group were used for the protein syntheses rate experiment, which ensured a statistical power at 1-β of 80% and α of 1%.
Data exclusions	No data were excluded.
Replication	For all experiments involved, 2-3 independent experiments were performed for for each data shown, and all attempts at replication were successful.
Randomization	The mice tissues, mouse and human cells for RNA sample collection were randomly chosen.
Blinding	Sample collection and RNA transfection are not blinded due to impracticality. Samples for RNA-seq were send out in single-blinded manner. Blinding were used during data analyses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Anti-Digoxigenin-AP Fab fragments (Roach, REF: 11093274910) used in the Northern blot experiments in Fig.2c-g, Fig. 3g-i, Fig.6k-r, Suppl Fig.2, and Suppl Fig.14 are diluted in 1:10000.
Validation	The polyclonal antibody from sheep is specific to digoxigenin and digoxin and shows no cross-reactivity with other steroids, such as human estrogens and androgens. This antibody is validated and extensively use in previous publications, the weblink below shows that this antibody has been used in 145 peer-reviewd papers: https://www.sigmaaldrich.com/catalog/product/roche/11093274910?lang=en&region=US

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The E14 mouse ESCs were kindly provided by Prof. Austin Smith (Stem Cell Institute, Cambridge, UK). The hESC line H9 was kindly provided by Prof. Ludovic Vallier (Stem Cell Institute, Cambridge, UK). The HeLa cell line is from ATCC (Cat# CCL-2).

Authentication	The self-renewal properties of mouse and human ESCs were confirmed by immunoflourescence marker expressions routinely and morphological characteristics. Mouse iPSCs were confirmed by using SSEA-1 surface staining and the endogenous Oct4-GFP reporter in transgene independent iPSCs (Dox off). The HeLa cell line were validated by ATCC using COI assay and STR analysis.
Mycoplasma contamination	All cell line were test negative for mycoplasma contamination in this study.
Commonly misidentified lines (See ICLAC register)	No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC and NCBI Biosample

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	9-10 weeks old C57BL/6J male mice were given access to food and water ad libitum and were maintained on a 12:12 h light-dark artificial lighting cycle, mice were housed in cages at a temperature of 22-25 °C, with 40-60% humidity.
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	Animal experiments were conducted under the protocol and approval of the Institutional Animal Care and Use Committee of the University of California, Riverside, the Institutional Animal Care and Use Committee of the University of Nevada, Reno, and the Animal Care and Use Committee of the Institute of Zoology, Chinese Academy of Sciences, China.

Note that full information on the approval of the study protocol must also be provided in the manuscript.